MEASURING THE DYNAMICS OF YOUNG AND SMALL BUSINESSES: INTEGRATING THE EMPLOYER AND NONEMPLOYER UNIVERSES

by

Steven J. Davis *
University of Chicago

John Haltiwanger *
University of Maryland

Ron S. Jarmin *
U.S. Bureau of the Census

C.J. Krizan *
U.S. Bureau of the Census

Javier Miranda *
U.S. Bureau of the Census

Alfred Nucci *
U.S. Bureau of the Census

and

Kristin Sandusky *
U.S. Bureau of the Census

CES 06-04 **February, 2006**

Abstract

We develop a preliminary version of an Integrated Longitudinal Business Database (ILBD) that combines administrative records and survey-based data for virtually all employer and nonemployer business units in the United States. In the process, we confront conceptual and practical issues that arise in measuring the importance and dynamic behavior of younger and smaller businesses. We also document some basic facts about younger and smaller businesses. In doing so, we exploit the ability of the ILBD to follow business transitions between employer and nonemployer status, and vice-versa. This aspect of the ILBD opens a new frontier for the study of business formation and the precursors to job creation in the U.S. economy.

JEL Codes: L22, M13

Keywords: Data integration, nonemployer businesses, LBD

* We gratefully acknowledge support from the Kauffman Foundation and the U.S. Census Bureau. For helpful comments on earlier drafts, we thank Mark Roberts, Tim Dunne, Thomas Holmes, Robert Strom, Robert Litan, E.J. Reedy, Ying Lowery, Richard Boden, Lucia Foster and the participants of the NBER/CRIW conference on Producer Dynamics in April 2005 and an AEA/SBA session in January 2005. We also would like to thank Paul Hanczaryk for helping us to understand the Census Bureau's nonemployer data better. This work has undergone a more limited review than official Census Bureau publications. The views, findings and opinions expressed in this work are those of the authors and do not necessarily indicate concurrence by the U.S. Census Bureau. All results have been reviewed to ensure confidentiality.

I. Introduction

The measurement of economic activity by federal statistical agencies focuses greater attention on larger, more mature business units. This data gathering strategy has two clear advantages. First, it yields greater accuracy in estimating the level of economic activity, whether "greater attention" takes the form of higher sampling probabilities or more careful auditing and editing. Second, it is easier to identify and promptly capture the activity of large, long-established business units. On both counts, the desire for a cost-effective approach to measuring the level of economic activity leads naturally to a focus on larger, more mature units.

There are, however, drawbacks to this data gathering strategy. When responses to shocks and new developments in the economy vary systematically with business size or age, a focus on larger and more mature units can yield less accurate, potentially misleading measures of changes in economic activity. As a simple example, consider the situation when younger and smaller business units are relatively sensitive to aggregate shocks. In this case, a cost-effective approach to estimating short-term growth rates can require the over sampling of younger and smaller business units, and there is tension between a sample design optimized for the level of activity and one optimized for the growth rate. More important, the traditional focus on larger and more mature units limits our ability to measure and study the early lifecycle dynamics of businesses and to evaluate theories of business formation, selection and growth.

This paper reports our initial efforts to remedy these drawbacks. We develop a preliminary version of an Integrated Longitudinal Business Database (ILBD) that combines administrative records and survey-based data for virtually all employer and nonemployer business units in the United States. In the process, we confront conceptual and practical issues that arise in measuring the importance and dynamic behavior of younger and smaller businesses. We also document some basic facts about younger and smaller businesses. In doing so, we exploit the ability of the ILBD to follow business transitions between employer and nonemployer status, and vice-versa. This aspect of the ILBD opens a new frontier for the study of business formation and the precursors to job creation in the U.S. economy.

There are 5.4 million businesses with employees in the United States as of 2000 and another 15.5 million with no employees. Most nonemployer business units are quite small, never become employers, and do not link to the employer universe by way of any ownership relation. Nonemployer businesses account for a modest four percent of aggregate U.S. business revenue within the year, but a substantial number of employers originate as nonemployer businesses.

Our analysis focuses on forty industries for which smaller and younger businesses play especially important roles. These industries account for nearly half of all nonemployer business units and thirty six percent of nonemployer revenues. Within these industries, nonemployers account for fourteen percent of business revenues. In addition, more than eleven percent of the employers in these industries are connected by some type of ownership link to the nonemployer business universe within the previous eight years. Some of these linkages reflect nonemployer businesses that become employers, but other linkage patterns arise frequently. There are many instances of employer and nonemployer business units that operate simultaneously under common ownership. In addition, many employers operate establishments that predate the

nonemployer businesses to which they link.

Over a three-year horizon, five percent of the roughly seven million nonemployers in our selected industries transit to the employer universe. Transiting businesses are relatively large compared to other nonemployers, and they grow considerably faster prior to transition to employer status. Mean revenue growth for transits in the year before transition is thirteen percentage points higher than for other nonemployers. Transits also play a nontrivial role in the formation of employer businesses. For example, they account for sixteen percent of young employers (zero to three years old) in our selected industries.[2]

As another step toward an integrated perspective on the dynamics of young and small businesses, we compare the growth patterns of employer and nonemployer businesses. We focus on revenue as an activity measure because it is available for all business units, employers and nonemployers alike. As in previous research (e.g., Davis and Haltiwanger, 1999), we find that the mean and dispersion of growth rates are higher for younger employers. The dispersion in growth rates declines sharply with age for nonemployers and, at any given age, is much greater for nonemployers than employers. Mean revenue growth rates are U-shaped in size among nonemployer. The dispersion in revenue growth rates declines with size for both employers and nonemployers.

The paper proceeds as follows. Section II discusses the construction of the ILBD and presents some basic facts about employer and nonemployer businesses. Section III investigates ownership linkages over time between employer and nonemployer businesses. We measure the incidence of ownership linkages between employer and nonemployer businesses, characterize some aspects of these linkages, and quantify business transits between nonemployer and employer status. Section IV documents revenue growth and dispersion patterns by business size and age. Section V discusses next steps in our research program, and Section VI offers concluding remarks.

II. Constructing an Integrated Longitudinal Business Database

A. Overview of Main Tasks and Previous Work

In terms of data development, our objective is to build a fully Integrated Longitudinal Business-level Database (ILBD) that covers all employer and nonemployer business units in the nonfarm private U.S. economy. We construct the initial version of the ILBD for the years 1992 and 1994-2000, and we plan updates for later years in future enhancements of the ILBD. Key data on nonemployers are unavailable for 1993.

From an analytical perspective, the presence or absence of employees is simply another business characteristic to be measured. From a database development perspective, however, integrating the Census Bureau's employer and nonemployer business universes is a major undertaking. The main tasks fall into three categories. One task is to integrate the employer and nonemployer universes on a year-by-year basis,

[2] As we explain below, these estimates are provisional in a number of respects. We believe they understate the role of transiting businesses, perhaps by a substantial amount, because they reflect a conservative algorithm for linking employer and nonemployer businesses.. Future enhancements to the ILBD will use more sophisticated and comprehensive matching algorithms.

ensuring that each unique business entity is counted once, and only once. A second task is to construct longitudinal links for business units within each universe. A third task is to construct contemporaneous and dynamic ownership links across universes between employer and nonemployer business units. To carry out these tasks, we build on previous work by Jarmin and Miranda (2003) to create the Longitudinal Business Database (LBD), which contains annual data from 1975 to 2001 for all nonfarm private employers. We also build on previous efforts to construct longitudinal links within the nonemployer universe by Nucci and Boden (2003) and Boden and Nucci (2004).

B. Source Data for the ILBD

Census Bureau business registers draw on payroll tax records, corporate and individual income tax returns, applications for an Employer Identification Number (EIN), and various Census Bureau business surveys. The data available to the Census Bureau depend on the legal and tax status of a business and, in certain respects, on the size of the business and the number of its locations. For large corporations, routine data inputs include payroll records and particular items from corporate income tax returns, augmented by direct Census Bureau collections for multi-location companies. For sole proprietors, partnerships and single-location corporations with employees, routine data inputs include payroll records, certain items from income tax returns and periodic Census Bureau surveys such as the quinquennial Economic Census. For nonemployer businesses, routine data inputs derive mainly from income tax returns. Table 2.1 lists the most important administrative and survey sources for key variables in the employer and nonemployer universes.

Linking records from different sources requires common business identifiers. Businesses with employees have unique Employer Identification Numbers (EINs) and other Census Bureau identifiers. Some nonemployer businesses also have an EIN, but most do not and instead are tracked by the person ID of the business owner, i.e., his or her Social Security Number (SSN).

To construct the ILBD, we must first ensure that administrative data from each universe are cleaned and ready for integration. On the employer side, this task has been largely accomplished in the work to create the LBD. The LBD contains high-quality longitudinal links for establishments, and for the most part it is easy to accurately determine establishment age.[3] Longitudinal establishment links are relatively straightforward to construct, because they are one to one, and because establishments typically have well-defined physical locations. Firms are more difficult to track over time, partly because firm-level links can be many to many. Work is underway at the Census Bureau to develop a rich set of longitudinal firm linkages. In the meantime, we define the age of a firm in the LBD as the age of its oldest establishment.

Longitudinal links are difficult to construct in the nonemployer universe for some of the same reasons that they are difficult to construct for firms in the employer universe. For example, when there is a change in the legal or tax status of a nonemployer business, its EIN or person ID can also change. (Person IDs do not change for individuals, but

[3]See Jarmin and Miranda (2003). The main outstanding issue with respect to the LBD concerns the delayed identification of new establishments owned by certain multi-unit companies. We are developing algorithms to retime these births. The retiming issue pertains only to the recognition date of establishment birth, not the company-wide level of revenues or other measures of economic activity.

ownership changes can yield a change in the person ID associated with a business.) In these instances, it is not straightforward to maintain longitudinal links for nonemployer businesses using data items that are routinely included in the Census Bureau's administrative records systems. On the employer side, direct Census Bureau collections provide this additional information, but there is no ready equivalent on the nonemployer side. Hence, in the face of changes in legal and tax status, longitudinal links for nonemployers require matching algorithms based on data items such as business name, location and industry.

Linking records across the employer and nonemployer business universes is further complicated by differences in the underlying administrative data and differences in the unit of analysis. The Census Bureau's Employer Business Register, which underlies the LBD, is a list of establishments (physical locations) maintained to serve as a mailing list for the Economic Census and as a frame for a variety of surveys. It relies heavily on administrative data and is augmented by direct Census Bureau collections.[4] Longitudinal linking is facilitated by establishment IDs (LBD Numbers and Permanent Plant Numbers or PPNs), EINs, enterprise IDs (Alphas), and business name and address information. In contrast, the Census Bureau's Nonemployer Business Register consists entirely of administrative data. The unit of analysis is a business entity, as recorded on a tax return. Our longitudinal links for nonemployer businesses currently exploit EINs, person IDs in the form of SSNs, business name information, and geographic information. One complication arises when an individual taxpayer reports income for multiple nonemployer businesses. For example, multiple Schedule C forms can be attached to a single 1040 tax form. We deal with these cases by aggregating to the level of a single tax filing, which is associated with a unique person ID or EIN. We then create longitudinal links across years using these numeric identifiers and business name. We describe our procedure for linking records across the business universes below.

C. Basic Facts about Employers and Nonemployers

Table 2.2 provides summary statistics for the employer and nonemployer business universes in 2000. There are about 15.5 million nonemployer businesses. Of these, 13.4 million are person ID units (sole proprietorships with no employees) and 2.1 million are EIN units (corporations, partnerships and other nonemployer business entities with EINs).[5] There are also about 5.4 million employer businesses. Of these, 182 thousand are multi-unit (MU) enterprises with multiple establishments, and the rest are single-unit (SU) businesses. While comparatively small in number, multi-unit enterprises account for sixty-one percent of aggregate U.S. business revenue. Nonemployer business units account for four percent of aggregate revenue, and single-unit employers account for thirty-five percent.

[4] In order to track the establishment structure of multi-unit enterprises, the Census Bureau conducts an annual Company Organization Survey. This survey covers all large multi-unit companies and a sample of smaller ones. During an economic census, all establishments of multi-unit companies receive survey forms.

[5] The distinction between person ID and EIN units can be complex. A sole proprietor with no payroll but positive receipts who has applied for an EIN can appear in both the person ID and EIN sections of the Nonemployer Business Register. That same proprietor can appear in the Employer Business Register as well. We assign all zero-payroll units to the nonemployer universe, even if they reside in the Employer Business Register.

Given the sheer size of the Census Bureau business registers and some complex issues of measurement, we focus on a selected set of forty industries for this paper. We choose industries with large numbers and relatively high revenue shares for young and small businesses. Dynamic links between employers and nonemployers are likely to be more important for these industries. We avoided industries with complex measurement issues related to financial holding companies, tax shelters, and special purpose financial entities. These aspects of corporate organizational structure are interesting, but they are not the focus of our data integration and analysis efforts.

Our analysis period overlaps with the transition from SIC to NAICS industry classifications, and the SIC-NAICS crosswalk is a many-to-many mapping. Our nonemployer data files contain 3-digit SIC codes prior to 1997 and 4-digit NAICS codes thereafter. The employer data files contain codes for both classifications from 1997 to 2000. Accordingly, we proceed as follows. For many exercises, we look backwards for businesses in selected 4-digit NAICS industries. For other exercises, we look forward from a year prior to the NAICS changeover at businesses in 3-digit SIC codes that correspond closely to our selected NAICS industries.

Table 2.3 provides summary information for our selected 4-digit NAICS industries. Legal Services has the largest number of employer businesses, almost 150 thousand. It also has the biggest employment and payroll with more than one million workers and more than 58 billion dollars in payroll. The highest-revenue industry for employers is Gasoline Stations at 187 billion dollars.[6] Other Personal Services has the largest number of nonemployer businesses, more than 800 thousand. The highest-revenue industry for nonemployers is Real Estate Agents and Brokers with almost 23 billion dollars.

Table 2.4 provides information about industry shares of aggregate business revenues and the relative size of the employer and nonemployer segments within industries. Nonemployer revenue shares range widely. At the upper end, nonemployers account for more than two-thirds of revenue in Independent Artists, Writers and Performers, and they account for at least thirty percent of revenues in twelve industries. At the lower end, nonemployers account for a mere twenty-four hundredths of a percent of revenue in Software Publishers, even though twenty percent of all businesses in this industry are nonemployers. Although extreme, the basic pattern in this industry is not uncommon; it reflects the enormously skewed size distribution of activity in many industries.

Figures 2.1 through 2.4 provide information about the age and size distributions of businesses in our selected industries. Age is measured in years since first appearance by a business entity in its respective business universe. For multi-location firms, business age equals the age of the firm's oldest establishment. We use revenue measures to compare size distributions across the two universes, because revenue is the only activity measure available for both universes.

As seen in Figure 2.1, older firms dominate economic activity in the employer universe. Firms that are at least eight years old account for almost seventy percent of employer revenues and more than forty percent of employer businesses in our selected industries. In contrast, older businesses play a much smaller role in the nonemployer universe. Firms that are at least eight years old account for only forty percent of

[6] High revenue in Gasoline Stations mainly reflects the cost of gasoline. The administrative data in the Census business registers typically does not include information on gross margins or material costs. The data are included in the Economic Censuses and various annual surveys.

nonemployer revenues. Very young nonemployers account for a larger share of business units and even revenue than, say, five-year old nonemployers.

As seen in Figure 2.2, most nonemployer businesses are quite small. Roughly a third of nonemployer businesses generate less than six thousand dollars in annual revenue. While large in numbers, these very small nonemployers account for a tiny fraction of business revenues. Somewhat larger business units account for much of the revenue generated by nonemployers. For example, nearly one-quarter of nonemployer revenue in our selected industries is generated by businesses with annual revenue in the range of 120 to 360 thousand dollars. In contrast, the size distribution of revenues has a very different shape in the employer universe. Almost seventy percent of employer revenue is generated by firms with more than three million dollars in annual revenue.

The tremendous variation in size across nonemployer and employer businesses exhibited in Figure 2.2 serves as a caution when drawing inferences about the behavior of "small and young" businesses per se. The wide size distribution reminds us that many nonemployer businesses are extremely small and often may represent a secondary or supplemental source of income to the household. Analyzing the dynamics of such businesses alongside much larger businesses is a challenge. In what follows, we often report results for both the share of business units and the share of revenue. The former provides more insights about the very small and more prevalent businesses, while the latter provides more insights into the contribution of larger businesses.

Figures 2.3 and 2.4 display the share of revenues and business units accounted for by young businesses (0-3 years old) and small businesses (less than $90,000 in annual revenue) in our selected industries. Figure 2.3 reveals wide variation across industries in the revenue and number shares of young and small businesses in the nonemployer universe. Figure 2.4 shows a similar pattern with respect to the revenue and numbers share of young businesses in the employer universe and with respect to the numbers share of small businesses. However, with the exception of Personal Care Services (NAICS 8121), businesses with less than $90,000 in annual revenue account for very small revenue shares, typically less than five percent.

D. The Matching Algorithm

In addition to the longitudinal links within each universe, we created a set of links between the employer and nonemployer universes for our selected industries. These links make use of numeric identifiers, business names, and other data items found on records in both universes. The numeric identifier for most nonemployer businesses is the owner's SSN. We have several numeric identifiers available in the Employer Business Register and from the LBD. The most useful for our purposes is the EIN, but many records in the Employer Business Register contain both an EIN and an SSN. For example, when a business owner applies for an EIN, he or she must fill out an SS-4 form for the IRS. This form includes the business name, the SSN of the business owner or chief officer and the EIN, all of which are included in the administrative records system in the Census Bureau business registers.

We take a conservative approach in matching records across the nonemployer and employer universes. In particular, we rely only on the EIN and SSN numeric identifiers and business name. Previous efforts to develop longitudinal identifiers for the LBD have shown that linkages can be significantly improved by the use of more sophisticated

probabilistic matching algorithms that exploit all of the relevant available information, and that take into account the reliability of the information. We plan to incorporate some of these techniques in future enhancements of the ILBD. For now, given the conservative nature of our algorithm, we understate the incidence of linkages between the two universes. For the same reason, we also understate the contribution of nonemployer businesses to the formation of employer businesses.

Although we use similar logic when linking the universes in either direction, it will be convenient to outline our method with reference to a particular cross section of employers. For example, consider all establishments with employees in our selected industries in the 2000 cross section. Using the longitudinal links in the LBD, we create a set of identifiers (EINs, SSNs and business names) associated with each establishment for each year back to 1992. We use these identifiers to determine matches to nonemployers from 1992 to 2000.

Since multiple establishments can have the same numeric and name identifiers, it is important to "unduplicate" the files by year and identifier before proceeding to match them to the nonemployer universe by EIN, SSN or name. In the employer universe, the most common form of duplication arises when a multi-location company operates many establishments under one EIN or business name. We perform the match using a file that contains only one record per unique value of the linking variables. We are most confident about cross-universe matches based on numeric identifiers. Therefore, we search for matches based on EIN, SSN and business name in that order. We experimented with a number of name-matching procedures, including the removal of vowels, various symbols, and commonly used abbreviations ("inc.", "ltd", etc.). However, after examining a sample of matches, we concluded that literal name strings produced highly reliable matches and appeared to miss very few "good" matches.

Given a set of matches between the two universes, we aggregate the LBD-based employer data within an industry to the firm level. The result is an LBD-based firm-level dataset with ownership linkages to nonemployer businesses and additional variables that describe the nature of the nonemployer records to which the employer firm links. At this point, the unit of observation is a business firm with at least one establishment operating in one of our selected industries. If a firm operates in more than one of our selected industries, we maintain separate records for each industry in which it operates.

III. Ownership Links and Transition Dynamics

A. Backward Links of Employers to Nonemployers

We turn now to ownership links between the employer and nonemployer business universes. We begin our analysis by considering all employer businesses in our selected industries in the 2000 cross section. Table 3.1 reports size and age distributions for these businesses in columns (2) and (3). These columns show familiar patterns by business age and size. That is, for a given cross section, the number of active businesses declines with age and size, but the bulk of activity – measured here by payroll – is concentrated in older and larger businesses.

The more novel elements of Table 3.1 appear in the last four columns, which provide

information about ownership links between the employer and nonemployer universes. The at-risk population for this analysis is all employer businesses in our selected industries with positive payroll in 2000. For these employers, we consider all ownership links to nonemployer businesses in the current and previous eight years, i.e., in 1992 or 1994 to 2000. Out of the 2.3 million employer businesses in our selected industries, about 266 thousand have ownership links to the nonemployer universe within the current or previous eight years based on the matching algorithm described above. Column (6) shows how these linked employers are distributed by employer size and age.

Columns (4) and (5) in Table 3.1 report the number and percentage of employers in each size class and age group with ownership links to nonemployers. Ownership links to the nonemployer universe occur for a sizable percentage of employers in every size class and age group. Among businesses with 1 to 4 employees, fourteen percent link to the nonemployer universe within the previous eight years. Among employers less than six years old, more than fifteen percent link to the nonemployer universe. The likelihood of ownership links to nonemployers declines after age five, but this pattern may simply reflect our inability to identify ownership links in 1993 and prior to 1992. The likelihood of links to the nonemployer universe is U-shaped in employer size. These results indicate that most employers do not have ownership links to the nonemployer universe, at least within the previous eight years. Nevertheless, the results provide strong evidence that ownership links to the nonemployer universe are a common phenomenon for employers of all sizes and ages. In this regard, recall that our results understate the incidence of ownership links between the two business universes because of our conservative matching algorithm, and because we cannot identify links in 1993 or prior to 1992.

Column (7) in Table 3.1 reports the pre-link distribution of nonemployer revenues for those nonemployers that link to the employer universe. That is, we sum deflated nonemployer revenue for each nonemployer record that links to the 2000 LBD for the year prior to the link. For example, if a 1997 nonemployer record links to an employer business that operates in 2000, we use the deflated 1996 nonemployer revenue in the computation[7]. Comparing columns (6) and (7) in the top panel yields the inference that relatively large nonemployers tend to link to larger employers. For example, employers with at least 500 workers account for 6.2 percent of pre-link revenues among linked nonemployers but only 0.24 percent of the linked nonemployers. Similarly, over 75 percent of the employer businesses with links to a nonemployer business have fewer than five employees, but they link to less than 48 percent of pre-link nonemployer revenues.

When thinking about the process of business formation and growth, we anticipate a pattern whereby some businesses start as nonemployers, grow over time, and eventually transition to employer status, perhaps continuing to grow thereafter. This pattern holds for many businesses in the ILBD, but it is certainly not the only linkage pattern that arises. This point is evident in the lower panel of Table 3.1, specifically in columns (6) and (7). More than thirty percent of nonemployer firms that link to the employer universe – and more than forty-five percent of pre-link revenues – involve links to employers that are at least eight years old as of 2000. All of these cases involve

[7] Not all linkages are simple transitions from the nonemployer universe to the employer universe. In some cases there can be multiple years where links exist between employer and nonemployer businesses that appear with positive revenue in each universe. Below we call these cases "duals." For the current exercise with use the nonemployer revenue in the year prior to the earliest link to avoid double counting.

nonemployer firms that link to previously established employer businesses. That is, they do not involve a nonemployer business that evolves into a new employer business.

Figure 3.1 highlights this point in greater detail by plotting the frequency distribution of the age difference between employers in the 2000 cross section and their linked nonemployers. Recall that we know the exact age in years for employers that first appear in the employer universe after 1975. For nonemployers, we construct an age measure based on first appearance in the Nonemployer Business Register. If a nonemployer business is present in the 1992 Nonemployer Register, we assign an age of eight years as of 2000. For each nonemployer that links to one of the employers in our selected industries in 2000, we then compute the difference between its age and the age of the employer to which it links.

According to Figure 3.1, sixty percent of nonemployers are older than the employer to which they link. These cases are consistent with the "standard" pattern whereby a nonemployer business evolves into a new employer business. The pronounced mode at a one-year age difference reflects businesses that transition to employer status one year after inception as a nonemployer[8]. Many other businesses operate in nonemployer mode for a few years before transitioning to employer status. In addition to these standard cases, Figure 3.1 shows a large number of linkages in which the employer business predates the nonemployer business. These "nonstandard" linkage cases reflect other types of ownership relations between the two business universes. For example, an individual who owns a business with employees may also generate consulting income in a nonemployer business. As another type of example, a corporate business with employees may establish nonemployer subsidiaries for legal, financial or tax reasons.

B. Forward Transitions for Nonemployers

To continue our exploration of linkages between the two business universes, we now conduct an analysis of forward-looking transitions. We first examine a population of 1994 nonemployer businesses and classify their operational status three years later in 1997. We carry out a parallel analysis below for a population of 1994 employer businesses.

Figure 3.2 summarizes three-year transition dynamics for the population of nonemployer businesses in one of our selected industries.[9] The at-risk population includes nonemployer businesses in 1994 that link to an employer business in 1994. We classify outcomes into four categories:

- **Exits** – businesses with positive revenue in the nonemployer universe in 1994, no revenue in the nonemployer universe in 1997, and no payroll in the employer universe in 1997.
- **Transits –** businesses with positive revenue in the nonemployer universe in 1994, no revenue in the nonemployer universe in 1997, but positive revenue

[8] This group deserves further study. A likely explanation for observing a nonemployer birth in year t and a transition to employer status in year t+1 is that the owner applied for an EIN in year t and did not hire employees until year t+1.

[9] The basic patterns for these transition dynamics are very similar over a six-year horizon, although the magnitudes change in the expected way, e.g., the share of activity accounted for by exits rises substantially.

in the employer universe in 1997.

- **Continuers –** businesses with positive revenue in the nonemployer universe in 1994 and 1997, and no revenue in the employer universe in 1997.
- **Duals –** businesses with positive revenue in the nonemployer universe in 1994, and positive revenue in both universes in 1997.

We compute the share of the 1994 nonemployer analysis population that falls into each category on an unweighted and revenue-weighted basis.[10]

Continuers account for sixty-two percent of nonemployer businesses and fifty-eight percent of their revenues in 1994. Exits account for another thirty-seven percent of nonemployer businesses and twenty-five percent of revenues. Transits account for only three percent of nonemployer businesses but seven percent of their 1994 revenues. While three percent is a small share of the population at risk, there are 7.4 million nonemployer businesses in our selected industries. In terms of raw numbers, about 220,000 transits from our 1994 population of nonemployers appear with positive payroll in the 1997 employer universe. This figure for transits amounts to about eleven percent of all employer firms in a given cross section of our selected industries. Figure 3.3 carries out the same type of transition analysis separately for SSN and EIN cases, and it shows that transits and exits occur with greater relative frequency among EIN cases.

Table 3.2 reports the percentage of young employers in 1997, and their revenues, accounted for by transits from the entire 1994 population of nonemployers. For each of our selected industries, we calculate transits to young employers in 1997 as a percentage of those young employers for both business units and payroll. Here, "young" refers to businesses with positive payroll in 1997 that first had positive payroll no earlier than 1994. Several industries stand out for a relatively big contribution of nonemployer transits to the population of young employers including Landscape and Horticultural Services, Painting and Paper Hanging, Carpentry and Floor Work, Insurance Agents and Brokers, Real Estate Agents and Managers, Computer and Data Processing Services, Automotive Repair Shops, Legal Services, Child Day Care Services, and Accounting, Auditing and Bookkeeping. The transition from nonemployer to employer status is relatively easy and common in these industries. On average, transits from the nonemployer universe account for sixteen percent of young employers in our selected industries and eleven percent of their revenues.

It is also interesting to compare the growth experiences of transits to other nonemployer businesses. As a first effort to address this issue, we calculate summary statistics on the distribution of revenue growth by later transition status. Our analysis population for these calculations consists of all nonemployers in our selected industries with positive revenue in 1996. Table 3.3 reports summary statistics on revenue growth

[10] Because the at-risk population includes nonemployers in 1994 with ownership links to employers in 1994, some of Duals reflect nonemployers with ownership links to the employer universe in 1994 and 1997. Similarly, some of the Transits reflect an individual who owns businesses in both universes in 1994 or business enterprises with entities in both universes in 1994 but only an employer business in 1997. In the next version of this paper, we will isolate businesses with "dual" status in 1994 and analyze them separately. We will also distinguish between nonemployer transits to new employers and nonemployer transits to previously established employers.

from 1995 to 1996 for this population, where the business-level growth rate measure is the change in annual revenue from $t-1$ to t divided by the simple average of revenue in $t-1$ and t.[11] Mean revenue growth for transits is considerably higher than for other nonemployers prior to transition to the employer universe. More surprising, the dispersion of growth rates for transits is also much lower prior to transition.

C. Forward Transitions for Employers

We now examine forward transitions for the 1994 population of employers in our selected industries. Figure 3.4 summarizes the three-year transition dynamics. As before, we group businesses in the at-risk population into four categories based on their status three years later. "Transits" now refers to businesses with positive payroll in 1994 and no payroll but positive revenues in 1997.

Continuers – businesses with positive payroll in both years and no ownership links to the nonemployer universe in 1997 – account for seventy percent of all employers and eighty percent of employer revenues. Exits account for twelve percent of employer revenues and twenty-two percent of employer businesses. The exit figures point to high death rates for employers in our selected industries, but they are considerably smaller than exit rates for nonemployers (Figure 3.2). Transits to the nonemployer universe account for only three percent of employers in 1994 and one and one-half percent of employer revenues. Since the number of employers is much smaller than the number of nonemployers, the number of such employers making a transition is relatively small compared to the transition from nonemployer to employer. Duals account for about five percent of employer revenues.

IV. Revenue Growth and Dispersion by Age and Size

The previous section investigates ownership linkages and transition dynamics between the employer and nonemployer business universes. We now investigate some basic aspects of business dynamics within each universe. There is a vast body of research on the relationship of business growth patterns to business size and age.[12] Most of this research restricts attention to businesses with employees, and often to a subset of such businesses that meet a minimum size threshold or that covers only publicly traded companies. The ILBD makes it possible to systematically analyze and compare the dynamics of employer and nonemployer businesses.

Figures 4.1 and 4.2 display the mean and dispersion of annual revenue growth rates by business age for employers and nonemployers. As before, the business-level growth rate is measured as the change in revenue from $t-1$ to t divided by the simple average of revenue in $t-1$ and t. To maximize the number of age categories for nonemployers, we focus on revenue growth from 1999 to 2000. For employer businesses, we also show the

[11] This measure yields growth rate outcomes in the closed interval [-2,2]. The measure is symmetric about zero and bounded, so that it affords an integrated treatment of entrants, exits and continuers. It underlies standard measures of gross job creation and destruction rates and related statistics. See Davis, Haltiwanger and Schuh (1996) for additional discussion.

[12] Dunne, Roberts and Samuelson (1989), Sutton (1997), Caves (1998) and Davis and Haltiwanger (1999) review various branches of this literature, which spans several decades.

average 1994-2000 patterns. Our measure of cross-sectional dispersion is the excess revenue reallocation rate: gross revenue gains at expanding units plus gross revenue losses at contracting units minus the absolute value of the net revenue change, all divided by aggregate revenue for the units under consideration. The excess reallocation rate is equivalent to the average absolute deviation of growth rates about zero, confirming its interpretation as a measure of dispersion in cross-sectional growth.[13] The measures in Figures 4.1 and 4.2 and elsewhere in this section are activity weighted, and they encompass entrants and exits as well as continuers. Given our measures, the mean growth rate for age-zero businesses is 200 percent, and their excess reallocation rate is zero.

Figure 4.1 shows that the mean growth rate of employers drops off very rapidly by age two and displays no clear relationship to age among older employers. The drop off in mean growth rates is even more rapid among nonemployers. Indeed, the mean growth rate for nonemployers is negative for ages greater than or equal to one. In this regard, it is important to recognize that we measure growth rates from a within-universe perspective in Figure 4.1 and the other figures in this section. For instance, the calculated growth rate is minus 200 percent for a nonemployer in 1999 that transits to the employer universe in 2000. In future work, we will exploit the ILBD to measure growth patterns for such businesses from an integrated perspective that captures their transformation to employer status.

Figure 4.2 shows that excess revenue reallocation rates tend to decline with business age, especially for nonemployers. Perhaps more important, the magnitude of excess revenue reallocation is very large for both types of businesses at all ages. The excess revenue reallocation rate for one-year old nonemployers exceeds eighty percent. Even for mature employers, excess revenue reallocation exceeds thirty percent in all age groups. These results underscore the tremendous amount of revenue expansion and contraction that takes place on a routine basis among U.S. business entities. In this respect, the results echo previous findings on the large magnitude of simultaneous job creation and destruction in Davis, Haltiwanger and Schuh (1996) and other work.

Figures 4.3 and 4.4 display revenue growth and excess reallocation rates by business size. The size categories are narrow at the lower end to reflect the revenue distribution among nonemployers. The mean growth rate is U-shaped with respect to size for nonemployers. Excess revenue reallocation rates decline sharply with size for employers and nonemployers. Excess reallocation is high for businesses of all sizes, exceeding twenty percent even for the largest businesses. Perhaps surprisingly, excess reallocation rates among businesses with less than $120,000 in annual revenue are greater for employers than nonemployers.

Figure 4.4 shows that dispersion in revenue growth declines with business size for both employers and nonemployers. For the very small businesses, dispersion is actually greater for employers, but once businesses have at least $180K in annual revenue, this relationship is reversed (i.e., nonemployers are more volatile). Interestingly, dispersion for employer and nonemployer businesses is about the same for the largest businesses.

As a robustness check, Figure 4.5 reports analogous measures by employer age using payroll rather than activity as a measure of activity. The payroll-based results are broadly

[13] See Davis and Haltiwanger (1999), who review the use of this measure in the literature on job flows.

similar to the earlier revenue-based results for employers, but the payroll measures appear to be less noisy. At least in part, the less noisy quality of the payroll-based results probably reflects the greater reliability of the business-level payroll data available to the Census Bureau for employer businesses. In contrast to the revenue-based results, the excess reallocation rate for payrolls declines monotonically with age.

V. Where Do We Go From Here?

The preceding sections describe the employer and nonemployer business universes, relate our efforts thus far to integrate the two universes, and present evidence on the dynamics of young and small businesses. In this section, we describe several challenges that we face in continuing the development of the ILBD.

A. Issues Common to Both Business Universes

One common issue is the conversion from SIC to NAICS industry codes following the 1997 economic censuses. A related but larger set of problems involves the reliability and accuracy of industry codes in the ILBD. Large portions of the two business universes rely almost exclusively on administrative records for source data on industry codes. As a rough generalization, the industry codes are less reliable and less precise for nonemployers and for smaller employers. Geographic identifiers also tend to be less accurate for nonemployers and smaller employers. In general, there are fewer sources of information for business-level records that derive entirely from administrative records, as compared to those that rely on administrative and survey sources.

Another common issue pertains to the interpretation of business revenue measures and their consistency over time. In particular, revenue measures can be affected by changes over time in income tax rules.

B. Issues in the Employer Universe

The quinquennial economic censuses gather a wealth of updated information on employer businesses. Much of this information identifies, for the first time, business-level changes that occurred in previous years. As a result, there is typically a large spike in the incidence of changes to industry and geography codes in census years. Similarly, there is a spike in newly identified establishments of multiunit firms. The spike reflects a delayed recognition of establishments opened by multiunit firms in previous years. Thus, for analyses that seek to measure the annual flow of new establishments, it is necessary to retime many of the spurious births that are first recorded in census years. We are currently developing algorithms for that purpose.

Multi-unit businesses above a size threshold are surveyed by the Census Bureau in the annual Company Organization Survey (COS). However, the list of such businesses is drawn from the prior economic census. These procedures mean that a firm's transition from single-unit to multi-unit status typically goes undetected until the next economic census. Further, new establishments operated by small multi-unit firms not covered by the COS are detected only at the economic censuses. In both cases, the economic activity measures for these new establishments are included with older establishments of the

parent company in the inter-censal years. Hence, the delayed recognition of some new establishments in inter-censal year leads to inaccurate business counts and, possibly, to an initially incorrect geographic and industrial classification for these new establishments.

Turning to another issue, the Census Bureau has made considerable progress in developing and maintaining longitudinal establishment identifiers for employer businesses, but the development of firm-level longitudinal identifiers remains an open area for research and development. Standard firm-level identifiers automatically change when a business undergoes certain types of reorganization such as a change in its legal form of organization or a merger. In the analysis above, we dealt with this issue by equating firm age to the age of the oldest establishment operated by the firm. However, we identified entry and exit of firms based on changes in standard firm-level identifiers in the ILBD.

C. Issues in the Nonemployer Universe

Some data issues unique to the nonemployer universe reflect the relatively recent availability and development of annual nonemployer files at the Census Bureau. For example, we cannot trace the inception of nonemployer businesses to years before 1992. In addition, as mentioned above, the nonemployer data rely very heavily on administrative sources. The construction of longitudinal links for nonemployer business units also raises several challenges, and our work in this area is at a relatively early stage of development. There is room for improving the longitudinal and cross-sectional linkages via name and address matching, the treatment of joint returns for proprietorships (where there are separate firm identifiers for the filer and his or her spouse), and the reliability of employer identification information for nonemployer proprietors.

D. Integrating the Two Business Universes

Bringing together the employer and nonemployer universes raises many additional issues. Refined matching techniques may yield a significant increase in the incidence of ownership links across the universes. In addition, we would like to better classify the ownership linkages between employer and nonemployer business units. For example, we would like to distinguish between nonemployer businesses that migrate to employer status within the year and nonemployer businesses that operate continuously, either as part of a larger enterprise with employees, or as a separate business entity under common ownership.

Studies of the integrated business-level data also face other challenges. First, the standards for classification by industry and geography differ between the two universes. These differences limit our ability to isolate narrowly defined industries and regions. Second, at the most basic level, the unit of observation differs between the two universes. For employers, the fundamental unit of observation is an establishment. For nonemployers, it is a tax return.

E. Integrating Employee Records with the ILBD

Another exciting direction for future research is the integration of *employee* data with the ILBD. Using the longitudinal matched employer-employee data from the LEHD program at the Census Bureau, demographic and earnings data for the universe of

employees can be integrated with the ILBD files.[14] Integration of the employer, nonemployer, and employee data will provide an unprecedented opportunity to study firm, job and worker dynamics. For example, it will be possible to follow someone who first works as an employee in a specific industry, then starts a small nonemployer business on the side, and later opens an employer business. More generally, for questions about where and how employer businesses originate, it will be very useful to know the demographic characteristics of business founders and their previous histories as employees and business owners.

VI. Concluding Remarks

It is tempting to think of the nonemployer business universe as a vast nursery for employer businesses. According to this view, many nonemployers evolve into employers and a few eventually grow into giant corporations that generate thousand of jobs. However, as our results confirm, most nonemployer business are quite small and never become employers. Indeed, it is misleading to think of all records in the nonemployer universe as "businesses" in the usual sense. Many nonemployer records reflect side jobs, hobby businesses or occasional consulting engagements that generate a little extra income for households that depend primarily on wages. One goal of our research is to help the Census Bureau develop algorithms that can distinguish hobby businesses, for example, from other types of nonemployer businesses, including entrepreneurial undertakings that might evolve into larger businesses with employees.

There is, however, a kernel of truth in the "vast nursery" view. Our evidence shows that migrants from the nonemployer universe account for a sizable share of young employer businesses in the industries we study. These migrants make up sixteen percent of young employers (zero to three years old) and account for eleven percent of their revenues. These figures probably understate the role of businesses that transition from nonemployer to employer status because of our conservative matching algorithms. In any event, the results indicate that a significant fraction of employer businesses originate as nonemployer businesses.

On the data front, this study takes important strides in developing an Integrated Longitudinal Business-level Database. Considerable work lies ahead, but the ILBD is already yielding useful information about the dynamics of young and small businesses. Major strengths of the LBD include comprehensive industry and geographic coverage, longitudinal links for establishments and firms, easy linkability to the large number of business surveys housed at Census and an integrated treatment of employer and nonemployer business. The ILBD makes it possible to examine the behavior over time of virtually all businesses in the U.S. economy, employers and nonemployers alike, with robust samples and even entire populations.

[14] Another obvious and important direction for future work is the integration of the Characteristics of Business Owners (CBO) and Survey of Business Owners (SBO) data sets into the ILBD. Holmes and Schmitz (1995), amongst others, have shown the rich analysis that can be conducted with the CBO.

References

Boden, Richard and Alfred Nucci, "Business Dynamics among the Smallest and Newest of Businesses: Preliminary Analyses of Longitudinally Matched Nonemployer and Business Register Data." Annual Meetings, Eastern Economic Association, Washington, D.C. March, 2004.

Caves, Richard E., "Industrial Organization and New Findings on the Turnover and Mobility of Firms" *Journal of Economic Literature*, 36 (4), 1998.

Davis, Steven J. and John Haltiwanger, "Gross Job Flows" in *Handbook of Labor Economics* (Orley Ashenfelter and David Card, eds.), North Holland, 1999.

Davis, Steven J., John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction*, MIT Press, 1996.

Dunne, Timothy, Mark Roberts and Larry Samuelson, "The Growth and Failure of U.S. Manufacturing Plants" *Quarterly Journal of Economics*, November 1989.

Holmes, Thomas and James Schmitz, "On the Turnover of Business Owners and Business Managers," *Journal of Political Economy*, October 1995.

Jarmin, Ron, and Javier Miranda, "The Longitudinal Business Database," CES Working Paper, 2003.

Nucci, Alfred and Richard Boden, "Demography of Nonemployer Businesses -- Preliminary Evidence from the United States." Paper Presented at the Comparative Analysis of Enterprise (micro) Data Conference, London, England, September 15-16, 2003

Sutton, John, "Gibrat's Legacy," *Journal of Economic Literature*, 35 (1), 1997.

Table 2.1 (a) Employer Business Register

Variable	Source Type	Source Details	Line Number
Name and Address	Survey	Physical address from Company Organization Survey, Annual Surveys, or Economic Censuses in Census years	
	Administrative	Beginning in 1998, physical address from form ss-4 for births; before 1998, mailing address from form ss-4	4a,4b; 5a,5b
		Beginning in 1998, physical address from form 941 for all businesses; before 1998, mailing address from form 941 for all businesses	Top of form
		Beginning in 1998, physical address from IRS income tax form for all businesses; before 1998, mailing address from IRS income tax form for all businesses	Top of form
Industry	Survey	Reported industry code from Company Organization Survey, Annual Surveys, or Economic Censuses in Census years Derived from the 1992 Economic Census--respondent reported classification Derived from a current survey (County Business Patterns (CBP), Company Organization Survey (COS)/Annual Survey of Manufactured (ASM), Current Industrial Report (CIR), Business Sample Revision (BSR) CBP Safeguard Review or intercensal refiles	
	Administrative	Derived from the Bureau of Labor Statistics	
		Derived from the Social Security Administration Derived from the Internal Revenue Service Principal Business Activity code, obtained from the (Business Master File	
Employment	Survey*	Company Organization Survey, Annual Surveys, or Economic Censuses in Census years	
	Administrative	IRS Form 941	Line 1
		IRS FICA wages	
		IRS total compensation	
		Imputed	
Payroll	Survey*	Company Organization Survey, Annual Surveys, or Economic Censuses in Census years	
	Administrative	IRS Form 941	Line 2
		IRS FICA wages	
		IRS total compensation	
		Imputed	
Revenue	Survey*	Company Organization Survey, Annual Surveys, or Economic Censuses in Census years	
	Imputed from EIN-level data	1120 – Gross receipts or sales less returns and allowances	Line 1c
		1120-A - Gross receipts or sales less returns and allowances	Line 1c
		1120F - Gross receipts or sales less returns and allowances	Section II, line 1a
		1120L - Gross income	Line 9
		1120-PC - Gross income	Sch A: line 14
		1120-RIC - Total income	Line 8
		1120S - Gross receipts or sales less returns and allowances	Line 1c
		1065 – Gross receipts or sales less returns and allowances	Line 1c
		990 - Total revenue	Line 12
		990-C - Gross receipts or sales less returns and allowances	Line 1c
		990EZ - Total revenue	Line 9
		990-PF - Total revenue	Line 12
		1040C - Gross receipts or sales less returns and allowances	Line 3

* = Multi-Unit Establishments Only

Table 2.1(b) Nonemployer Business Register, Proprietors - SSN Records

Variable	Source from IRS Form 1040, Schedule C
Name	Line C
Mailing Address	Line E
Legal Form of Organization	Implied by filing of IRS form 1040 Schedule C
Industry Code	Line B
Revenue	Line 3: Gross receipts or sales less returns and allowances

NOTE: All data for sole proprietors (including proprietorships jointly operated by husband and wife) are obtained form IRS form 1040, Schedule C.
NOTE: Line B of Schedule C reports the "Principal business or profession, including product or service." Based on this information, the IRS codes the Principal Business Activity (PBA) of the proprietorship. Based on the PBA code, the Census assigns a Tabulated Kind of Business (TKB) code. If the PBA is not reported on the tax form, then the Census uses the historic TKB code, if available.

Table 2.1 (c) Nonemployer Business Register, Partnerships and Corporations – EIN Records

Form Number	Form Description	Industry Code Source	Revenue Source
Form 1065	U.S. Partnership Return of Income	Line A (Principal Business Activity)	Line 1c: Gross receipts or sales less returns and allowances
Form 1120	U.S. Corporation Income Tax Return	Schedule K, Line 2a (Business Activity Code)	Line 1c: Gross receipts or sales less returns and allowances
Form 1120A	U.S. Corporation Short Form Tax Return	Part 2, line 1a (Business Activity Code)	Line 1c: Gross receipts or sales less returns and allowances
Form 1120 S	U.S. Income Tax Return for an S Corporation	Line B (Business Code)	Line 1c: Gross receipts or sales less returns and allowances
Form 1120 F	U.S. Income Tax Return of a Foreign Corporation	Line F1 (Business Activity Code)	Section II, 1c: Gross receipts or sales less returns and allowances
Form 1120 PC	U.S. Casualty and Property Insurance Company income Tax Return	Schedule I, Line 2	Schedule A: line 14 (gross income)
Form 1120 L	U.S. Life Insurance Company Income Tax Return	Schedule M, 2a-c (Kind of Company, Principal Business)	9 Gross Income
Form 1120 RIC	U.S. Income Tax Return for Regulated Investment Companies	Inferred by Form Type	8: Total Income
Form 1120 REIT	U.S. Income Tax Return for Real Estate Investment Trusts	Inferred by Form Type	8: Total Income

NOTE: All data for these businesses are derived from IRS income tax returns filed by the businesses.

NOTE: Name and Mailing Address taken from top of forms.

NOTE: Industry codes on Census Nonemployer database are IRS PBA codes obtained from the sources noted above and then converted to Census TKB codes. If the PBA is not available from the tax form, then the Census use the historic TKB code, if available.

NOTE: Legal form of organization implied by type of form submitted: 1065 filers are partnerships, all others are corporations.

Table 2.2 Summary Statistics for the Employer and Nonemployer Business Universes, 2000

		Nonemployer Business Universe				Employer Business Universe			
			40 Selected Industries				40 Selected Industries		
		All Industries	Number or Revenue	As Percent of All		All Industries	Number or Revenue	As Percent of All
# of Units (millions)	SSN Units	13.38	6.84	51	Singles	5.26	1.9	36
	EIN Units	2.15	0.54	25	Multi-Unit	0.18	0.06	31
	All	15.54	7.38	48	All	5.44	1.96	36
Revenue (billions $)	SSN Units	459.53	199.87	43	Singles	6,113.43	877.92	14
	EIN Units	251.74	55.58	22	Multi-Unit	10,758.04	664.52	6
	All	711.26	255.45	36	All	16,871.47	1,542.44	9

	Integrated Business Universe							
	All Industries				40 Selected Industries			
	Nonemployer Businesses		*Employer Businesses*		*Nonemployer Businesses*		*Employer Businesses*	
	EIN Units	SSN Units	Singles	Multi-Unit	EIN Units	SSN Units	Singles	Multi-Unit
Percent of Aggregate Revenue	1.43	2.61	34.77	61.19	3.09	11.12	48.83	36.96

Table 2.3 Summary Statistics for Selected Industries in the Integrated Business Universe, 2000

	Employers				Nonemployers	
Industry Description	Firms	Payroll	Employment	Revenue	Nonemployer Entities	Revenue
Animal production support activities	1.5	156	7	664	37.7	1,462
Painting & wall covering contractors	36.4	6,274	226	17,469	213.5	7,443
Carpentry & floor contractors	52.3	9,637	383	35,398	389.6	16,722
Roofing, siding, & sheet metal contractors	27.6	8,143	264	27,315	86.5	5,047
Concrete contractors	24.5	5,092	178	18,185	42.6	2,567
Printing & related support activities	23.2	14,226	420	41,613	26.7	1,486
Ship & boat building	1.1	3,788	105	16,896	0.4	47
Gasoline stations	52.2	12,282	837	187,841	9.4	1,682
Book, periodical & music stores	9.4	2,713	222	12,577	28.9	1,008
Florists	20.7	1,645	123	6,417	22.7	869
Taxi & limousine service	5.1	1,206	65	3,451	117.6	3,419
Couriers	2.3	15,654	550	18,610	1.1	111
Software publishers	6.7	23,009	249	49,988	1.7	120
Agencies & other insurance related activities	101.7	30,448	745	90,461	308.2	14,849
Offices of real estate agents & brokers	50.0	11,504	297	46,826	476.6	22,952
Activities related to real estate	39.5	15,052	484	38,058	356.3	18,274
Consumer goods rental	13.7	3,483	236	11,391	16.0	768
Legal services	147.7	58,514	1,055	149,400	206.3	11,626
Accounting, tax prep, bookkeep, payroll service	76.9	21,273	765	42,205	294.5	6,059
Computer systems design & related services	74.9	76,674	1,194	141,900	249.4	9,688
Management, sci & tech consulting services	76.5	32,718	729	67,277	355.2	16,796
Travel arrangement & reservation services	21.3	8,713	292	45,546	31.8	1,776
Services to buildings & dwellings	115.2	24,903	1,407	59,490	538.9	11,294
Offices of physicians	120.5	46,346	1,043	102,651	149.6	11,664
Offices of dentists	35.0	7,268	273	19,865	29.4	1,907
Offices of other health practitioners	74.7	8,814	346	31,356	235.2	9,053
Individual & family services	21.1	9,411	478	11,919	65.3	1,096
Child day care services	39.2	9,747	691	14,125	516.6	6,263
Agents, managers for artists & public figures	1.8	502	10	1,549	25.0	867
Independent artists, writers & performers	8.9	2,272	38	4,461	465.1	9,631
Rooming & boarding houses	1.5	182	13	717	9.7	281
Full-service restaurants	91.9	20,908	1,863	47,763	29.0	3,308
Limited-service eating places	94.4	17,164	1,743	53,707	36.8	3,086
Special food services	11.2	8,314	563	17,466	68.5	2,014
Drinking places (alcoholic beverages)	44.3	3,753	360	12,468	21.3	1,466
Automotive repair & maintenance	94.8	12,533	583	46,945	251.2	11,570
Personal & household goods R&M	22.4	2,337	98	8,685	247.9	6,147
Personal care services	59.8	4,109	303	9,770	552.4	11,776
Drycleaning & laundry services	30.8	6,613	373	17,731	33.0	1,650
Other personal services	16.0	3,688	173	12,288	835.8	17,609
Total Economy	5,443.40	3,773,003	113,658	16,871,471	15,536.07	711,264

Firms and employment in thousands. Payroll and Revenue in millions.

Table 2.4 Industry Revenue Shares and Business-Type Shares Within Industries, 2000

NAICS Code	Industry Description	Industry Revenues, Percent of Aggregate Business Revenue	Industry Revenues, Percent Accounted for by Employers	Percent of Business Entities in Industry that Are Nonemployers
1152	Animal production support activities	0.01	31	96
2352	Painting & wall covering contractors	0.14	70	85
2355	Carpentry & floor contractors	0.30	68	88
2356	Roofing, siding, & sheet metal contractors	0.18	84	76
2357	Concrete contractors	0.12	88	63
3231	Printing & related support activities	0.25	97	54
3366	Ship & boat building	0.10	100	26
4471	Gasoline stations	1.08	99	15
4512	Book, periodical & music stores	0.08	93	75
4531	Florists	0.04	88	52
4853	Taxi & limousine service	0.04	50	96
4921	Couriers	0.11	99	32
5112	Software publishers	0.28	100	20
5242	Agencies & other insurance related activities	0.60	86	75
5312	Offices of real estate agents & brokers	0.40	67	91
5313	Activities related to real estate	0.32	68	90
5322	Consumer goods rental	0.07	94	54
5411	Legal services	0.92	93	58
5412	Accounting, tax prep, bookkeeping, payroll service	0.27	87	79
5415	Computer systems design & related services	0.86	94	77
5416	Management, scientific & tech consulting services	0.48	80	82
5615	Travel arrangement & reservation services	0.27	96	60
5617	Services to buildings & dwellings	0.40	84	82
6211	Offices of physicians	0.65	90	55
6212	Offices of dentists	0.12	91	46
6213	Offices of other health practitioners	0.23	78	76
6241	Individual & family services	0.07	92	76
6244	Child day care services	0.12	69	93
7114	Agents, managers for artists & public figures	0.01	64	93
7115	Independent artists, writers & performers	0.08	32	98
7213	Rooming & boarding houses	0.01	72	86
7221	Full-service restaurants	0.29	94	24
7222	Limited-service eating places	0.32	95	28
7223	Special food services	0.11	90	86
7224	Drinking places (alcoholic beverages)	0.08	89	32
8111	Automotive repair & maintenance	0.33	80	73
8114	Personal & household goods R&M	0.08	59	92
8121	Personal care services	0.12	45	90
8123	Drycleaning & laundry services	0.11	91	52
8129	Other personal services	0.17	41	98

Table 3.1 Employer Links to Nonemployers by Size and Age of Employer, Employers in Selected Industries in 2000

(1) Firm Size in 2000, Number of Employees	(2) Number of firms in size class	(3) Percent of payroll in size class	(4) Number of firms with backward links to nonemployers	(5) Percent of firms in size class with backward links	(6) Percent of all backward links to nonemployers	(7) Percent of non-employer pre-link revenues
a)1-4	1416292	9.81	200252	14.14	75.36	46.73
b)5-9	432027	9.02	34590	8.01	13.02	13.09
c)10-19	242636	10.58	16656	6.86	6.27	12.38
d)20-49	139368	13.43	9229	6.62	3.47	8.62
e) 50-99	36886	8.56	2526	6.85	0.95	6.18
f)100-249	16426	8.83	1362	8.29	0.51	4.48
g)250-499	4332	7.16	489	11.29	0.18	2.32
h) 500+	3103	32.62	634	20.43	0.24	6.20
Total	2291070	100.00	265738	11.60	100.00	100.00

(1) Firm Age in 2000, Years Since First Appearance in Employer Universe	(2) Number of firms in age group	(3) Percent of payroll in age group	(4) Number of firms with backward links to nonemployers	(5) Percent of firms in age group with backward links	(6) Percent of all backward links to nonemployers	(7) Percent of non-employer pre-link revenues
0	213292	2.28	35082	16.45	13.20	11.01
1	190690	3.27	31314	16.42	11.78	9.08
2	170091	3.40	27197	15.99	10.23	7.95
3	153400	3.44	25244	16.46	9.50	6.46
4	134315	3.40	20675	15.39	7.78	5.49
5	117723	3.10	20615	17.51	7.76	4.48
6 – 7	199523	5.61	24823	12.44	9.34	10.81
8+	1112036	75.50	80788	7.26	30.40	44.74
Total	2291070	100.00	265738	11.60	100.00	100.00

Note: Column (7) reports the pre-link percentage distribution of nonemployer revenues for nonemployers that link to firms in the employer universe. To calculate this distribution, we first express nonemployer revenues in 2000 dollars using the GDP deflator for all goods and services. Then, for each nonemployer that links to the employer universe, we take the value of its deflated revenue in the year prior to its first link to a firm in the employer universe. We sum these values across all nonemployers that link to the 2000 LBD for our selected industries. The percentages reported in the table are based on this total value of "year prior to link" nonemployer revenue.

Table 3.2. Transits from the 1994 Nonemployer Population as a Percentage of Young Employer Businesses in 1997, Selected Industries

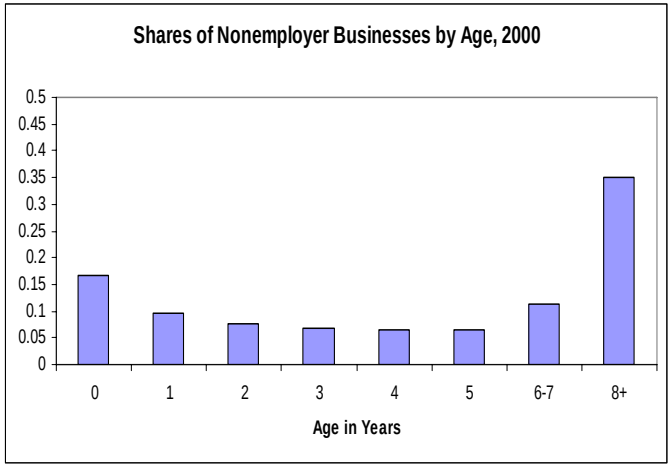
		Transits from 1994 Nonemployers to Young Employers in 1997 As a Percentage of Young Employers (0 to 3 Years Old)	
SIC Code	Industry Description	Percentage of Businesses	Percentage of Revenues
076	Farm labor and management services	26.49	22.73
078	Landscape and horticultural services	26.41	19.64
172	Painting and paper hanging	24.50	18.14
175	Carpentry and floor work	23.58	17.03
176	Roofing, siding, and sheet metal work	19.56	13.35
275	Commercial printing	13.57	7.44
367	Electronic components and accessories	9.71	3.99
412	Taxicabs	15.84	18.31
421	Trucking and courier services, except air	20.87	18.18
472	Passenger transportation arrangement	14.94	12.19
554	Gasoline service stations	8.59	6.62
581	Eating and drinking places	8.70	5.71
621	Security brokers and dealers	18.04	9.66
641	Insurance agents, brokers, and service	25.68	16.33
653	Real estate agents and managers	22.88	20.61
721	Laundry, cleaning, and garment services	18.25	13.13
723	Beauty shops	19.92	15.85
729	Miscellaneous personal services	22.82	24.59
737	Computer and data processing services	10.00	8.25
738	Miscellaneous business services	16.75	12.22
753	Automotive repair shops	19.47	12.23
784	Video tape rental	16.37	11.18
792	Producers, orchestras, entertainers	15.81	9.95
799	Misc. amusement, recreation services	12.53	8.44
801	Offices and clinics of medical doctors	11.65	8.19
802	Offices and clinics of dentists	17.55	14.09
803	Offices of osteopathic physicians	14.33	10.69
804	Offices of other health practitioners	18.73	13.19
811	Legal services	18.36	11.72
832	Individual and family services	10.90	6.03
835	Child day care services	22.31	13.06
872	Accounting, auditing, and bookkeeping	22.07	12.74
874	Management and public relations	12.41	7.79
	All Selected Industries	16.40	10.83

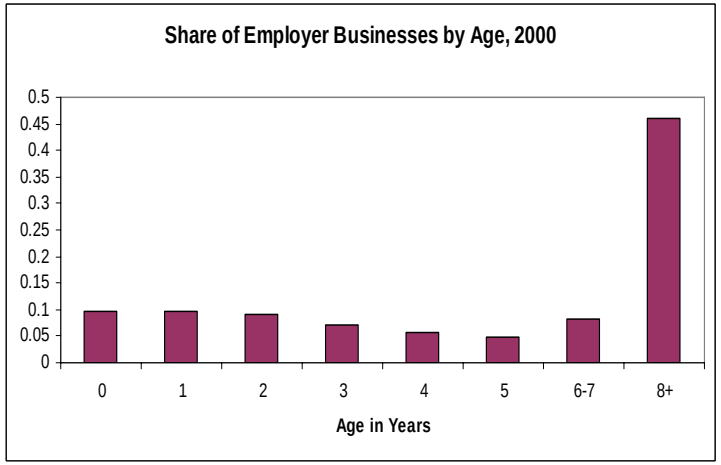
Table 3.3 Summary Statistics for the Distribution of Nonemployer Revenue Growth by Transition Status

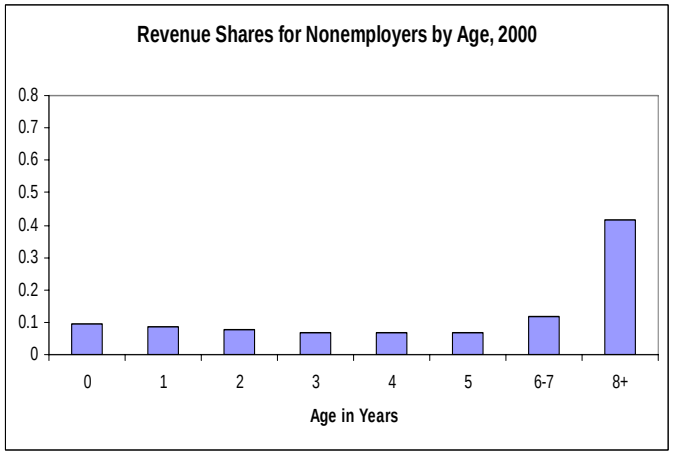
	Time Interval for Revenue Change	Transits from the Nonemployer Universe in 1996 to the Employer Universe in 1997	All Other Nonemployers with Positive Revenue in 1996
Mean	1995-96	0.024	-0.103
Median	1995-96	0.028	-0.011
90-10 Differential	1995-96	0.894	1.463

Notes: The analysis population contains all nonemployers in our selected industries with positive revenue in 1996. Table entries report summary statistics for the distribution of annual revenue growth rates from 1995 to 1996. The revenue growth rate is measured as the change in annual revenue from $t-1$ to t divided by the simple average of revenue in $t-1$ and t. All statistics are computed on a revenue-weighted basis. The table shows that transits from the nonemployer universe to the employer universe have a higher average growth rate and less dispersion in their growth rates in the year prior to transition.

Figure 2.1 Age Distribution of Business Numbers and Revenues within Each Universe







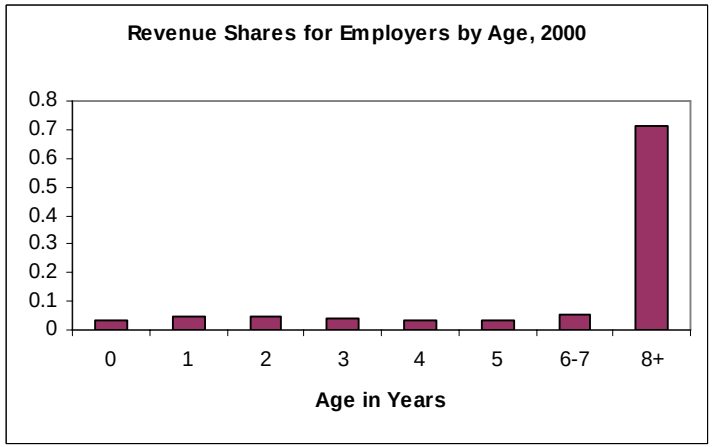
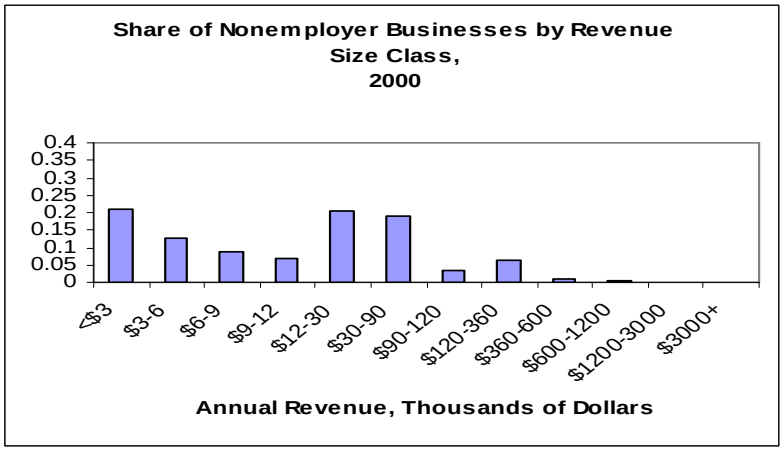
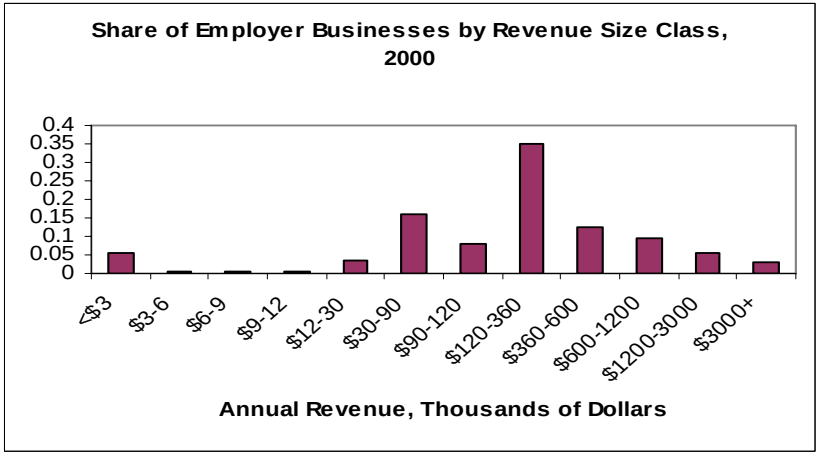
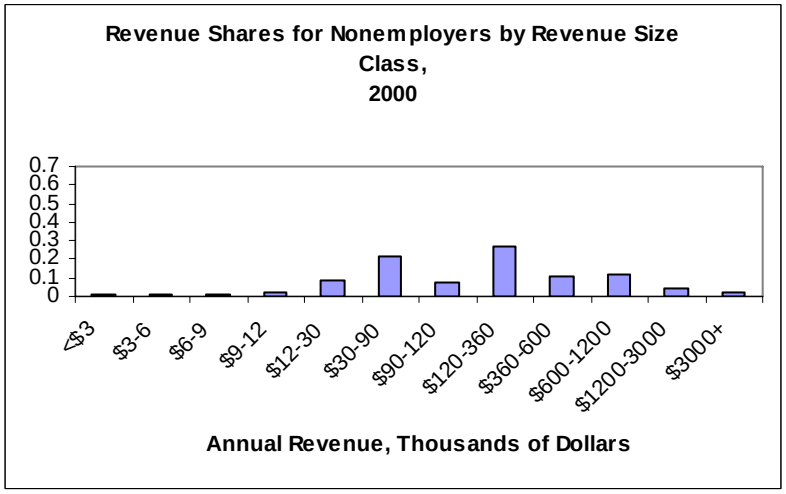


Figure 2.2 Size Distribution of Business Numbers and Revenues with Each Universe







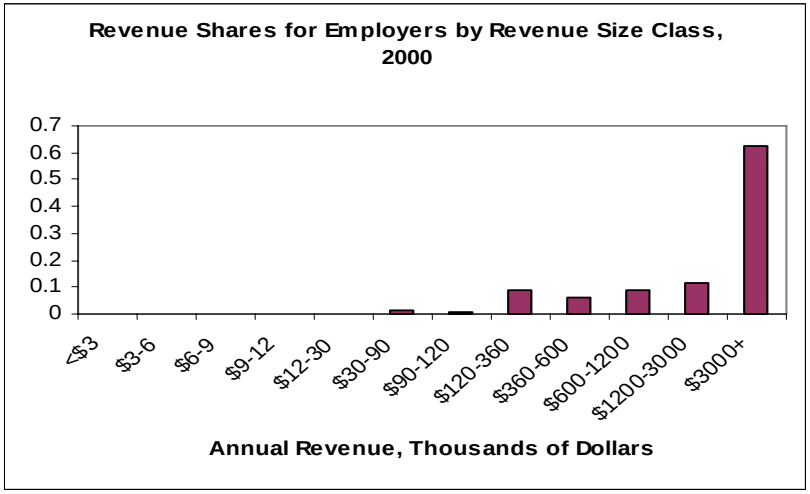


Figure 2.3 Share of Revenues and Business Numbers Accounted for by Young Firms, 0 to 3 Years Old, in Selected 4-Digit NAICS Industries

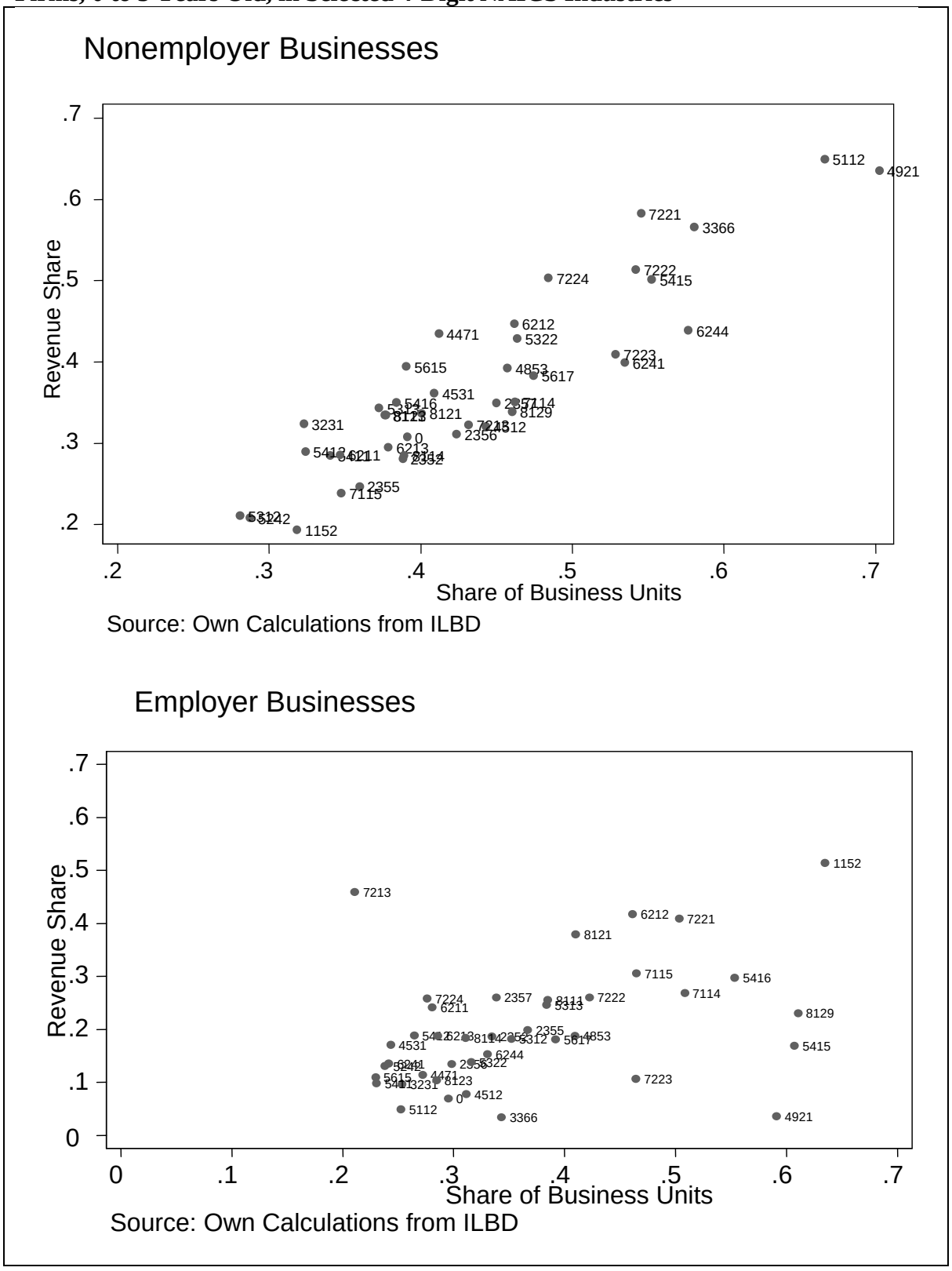


Figure 2.4 Share of Revenues and Business Numbers Accounted for by Small Firms, Less than $90,000 Annual Revenue, in Selected 4-Digit NAICS Industries

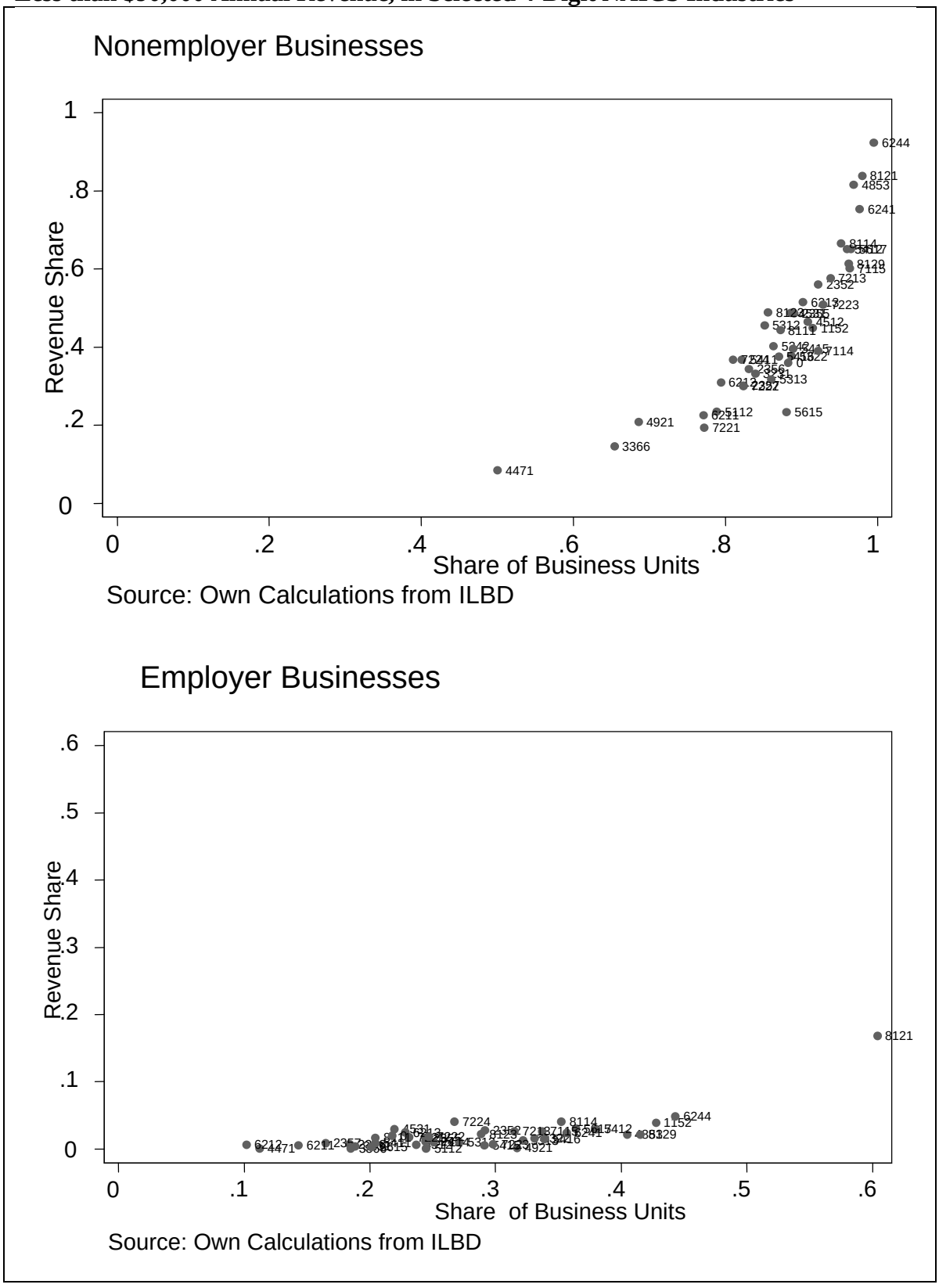


Figure 3.1

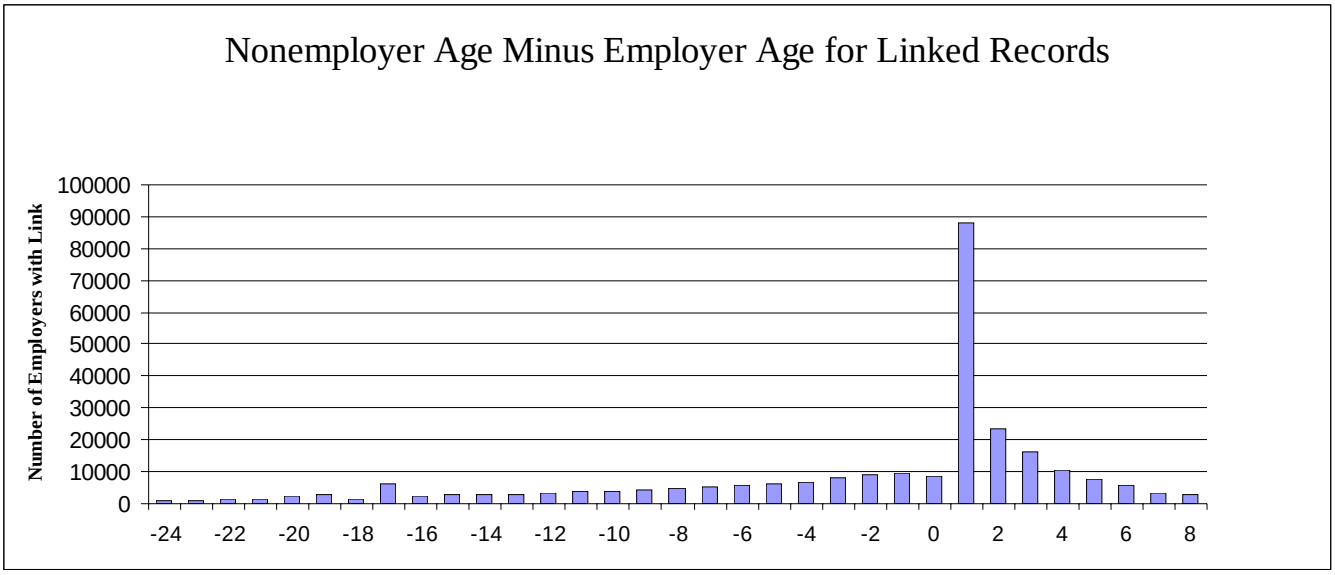

Nonemployer Age Minus Employer Age for Linked Records
Number of Employers with Link
100000
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
-24
-22
-20
-18
-16
-14
-12
-10
-8
-6
-4
-2
0
2
4
6
8

Figure 3.2

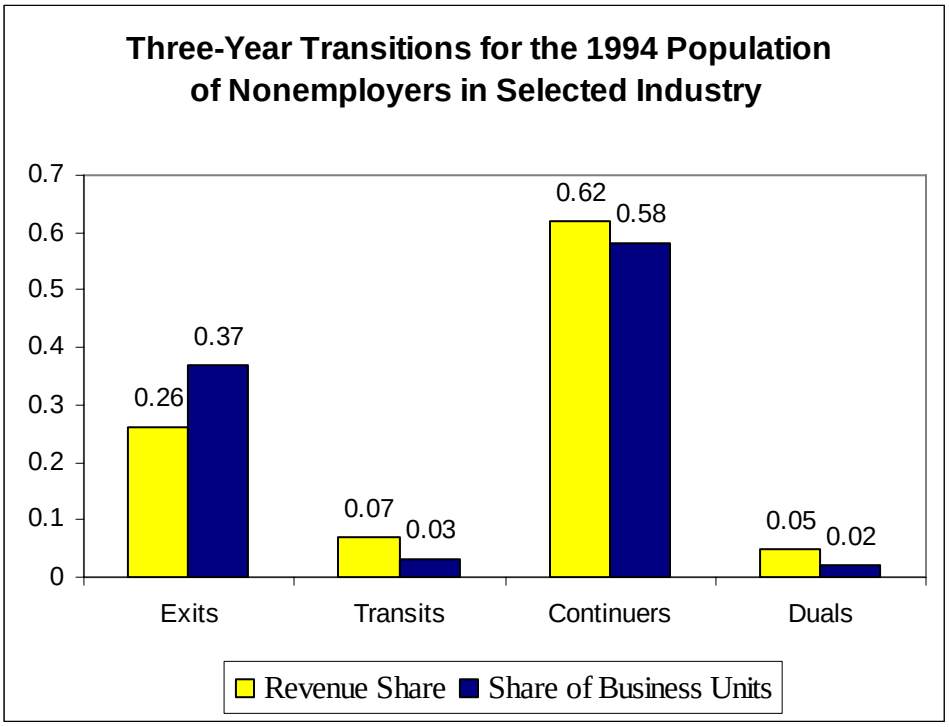
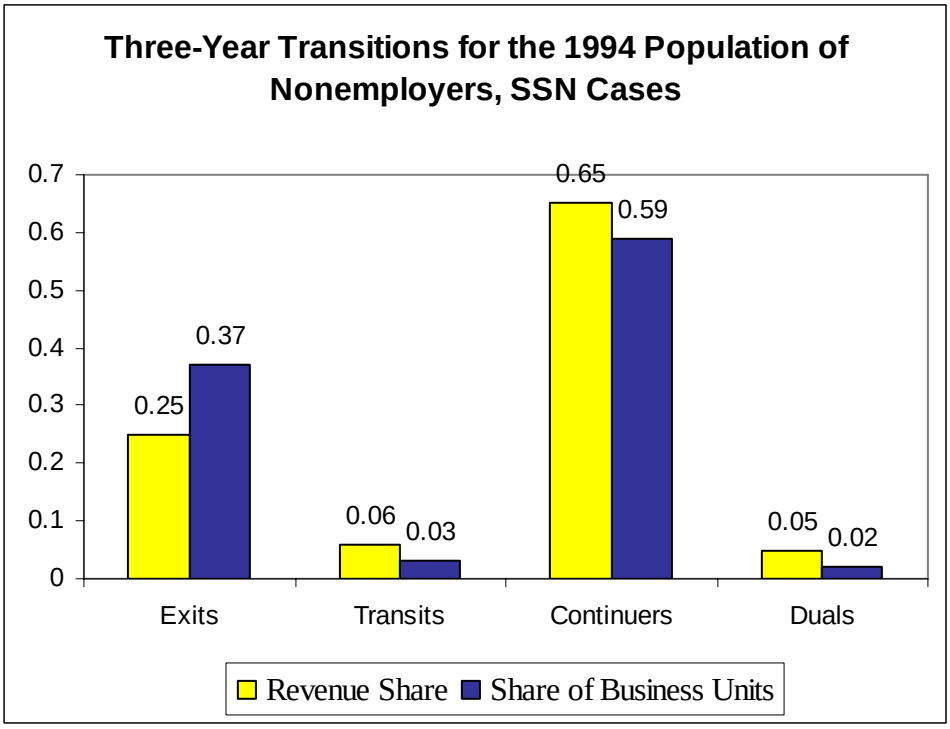

Three-Year Transitions for the 1994 Population of Nonemployers in Selected Industry
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.26
0.37
0.07
0.03
0.62
0.58
0.05
0.02
Exits
Transits
Continuers
Duals
Revenue Share
Share of Business Units

Figure 3.3


Three-Year Transitions for the 1994 Population of Nonemployers, SSN Cases
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.25
0.37
0.06
0.03
0.65
0.59
0.05
0.02
Exits
Transits
Continuers
Duals
Revenue Share
Share of Business Units

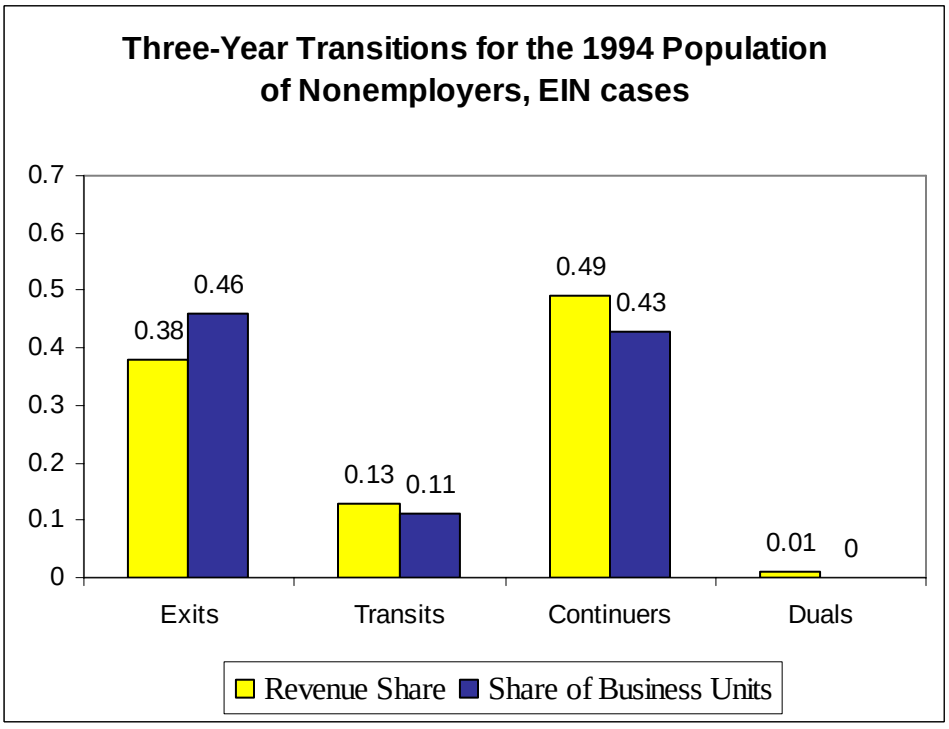

Three-Year Transitions for the 1994 Population of Nonemployers, EIN cases
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0.38
0.46
0.13
0.11
0.49
0.43
0.01
0
Exits
Transits
Continuers
Duals
Revenue Share
Share of Business Units

Figure 3.4

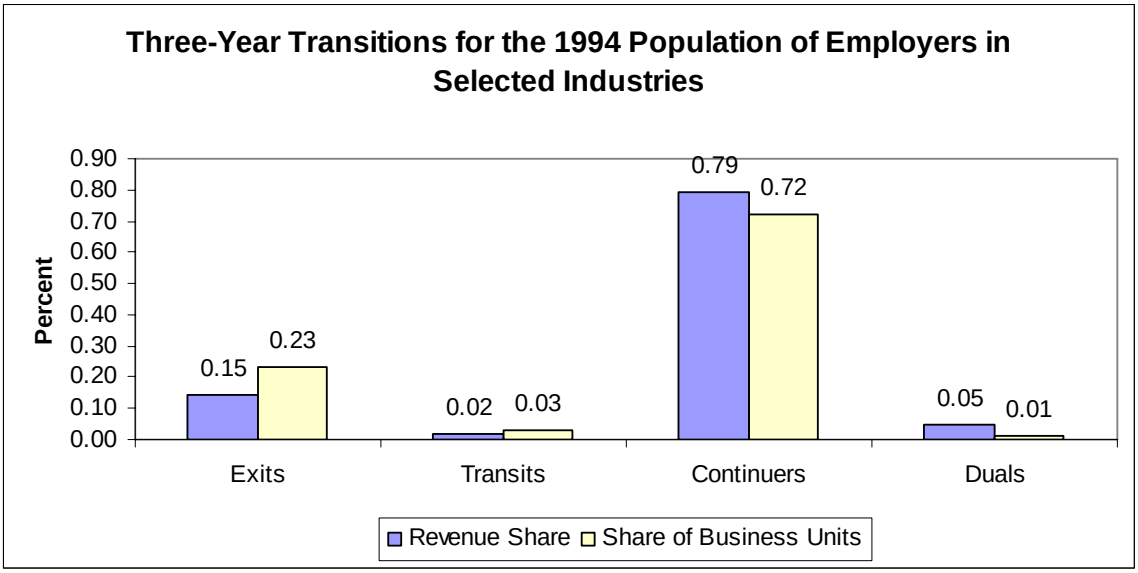

Three-Year Transitions for the 1994 Population of Employers in Selected Industries
Percent
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.15
0.23
0.02
0.03
0.79
0.72
0.05
0.01
Exits
Transits
Continuers
Duals
Revenue Share
Share of Business Units

Figure 4.1 Annual Revenue Growth by Business Type and Age

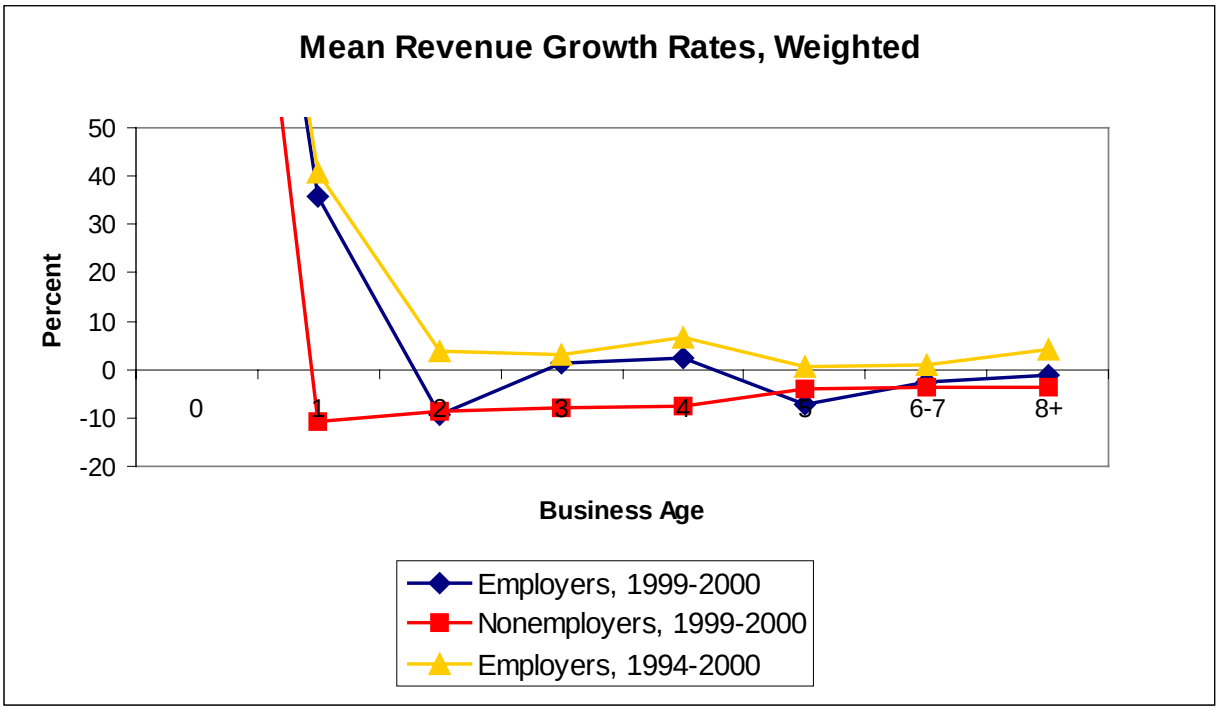


Figure 4.2 Excess Revenue Reallocation by Business Type and Age

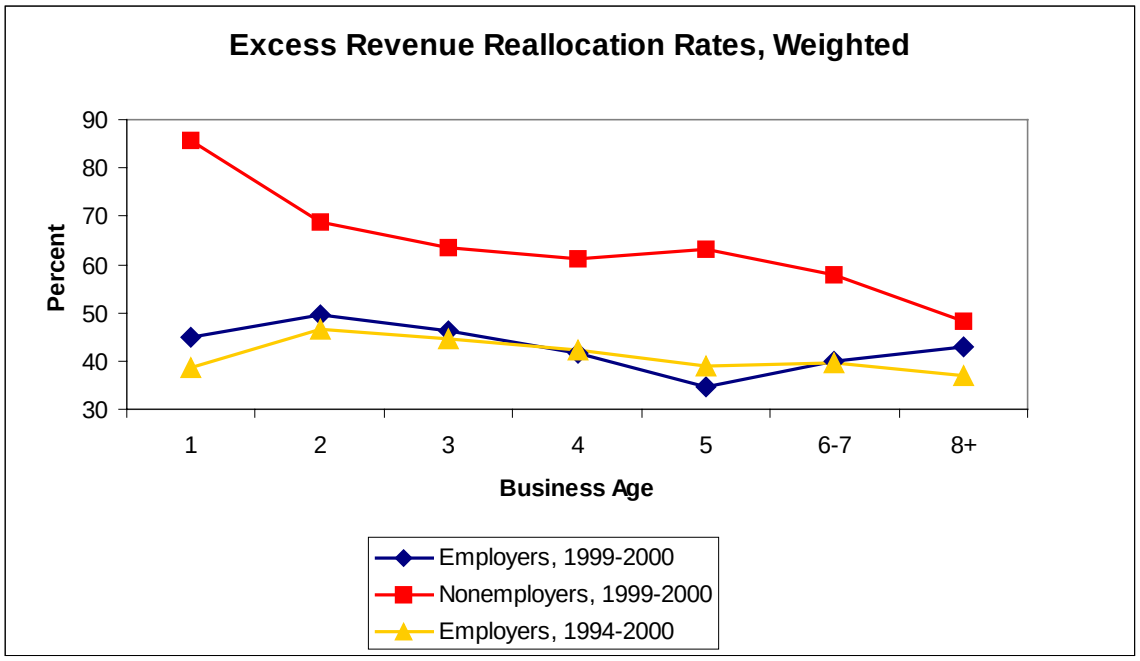


Figure 4.3 Mean Revenue Growth by Business Type and Size

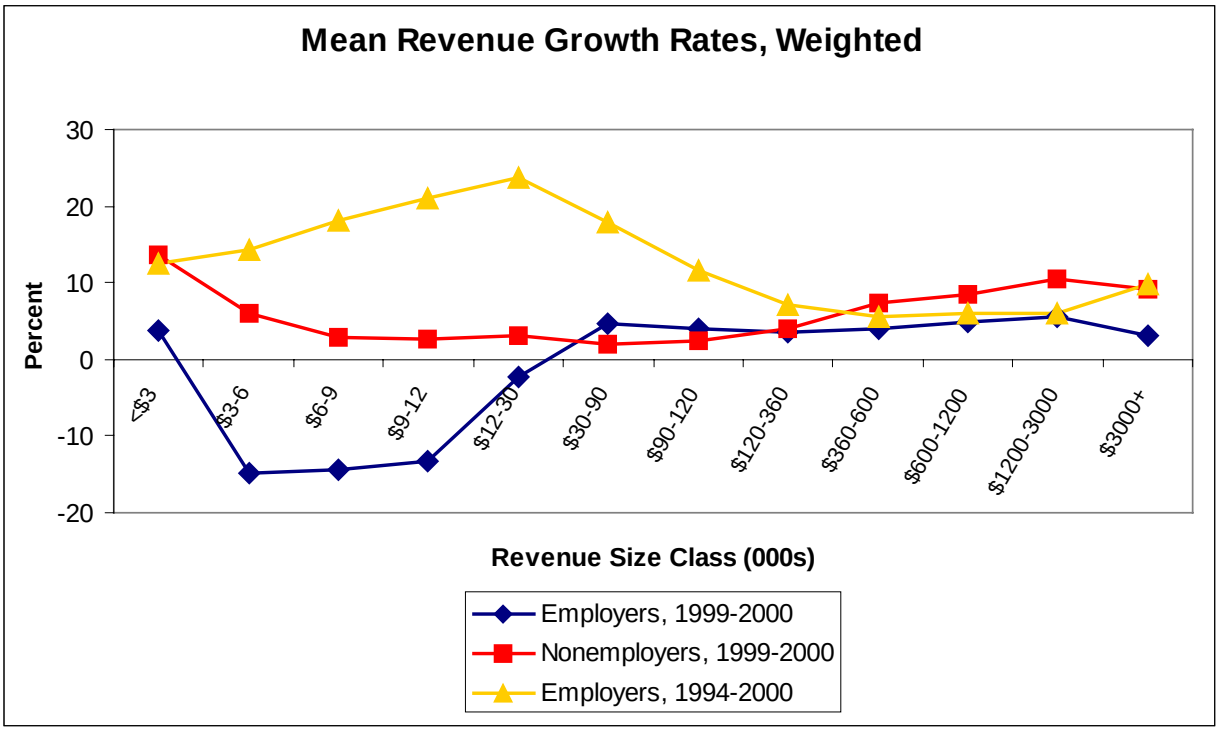


Figure 4.4 Excess Revenue Reallocation by Business Type and Size

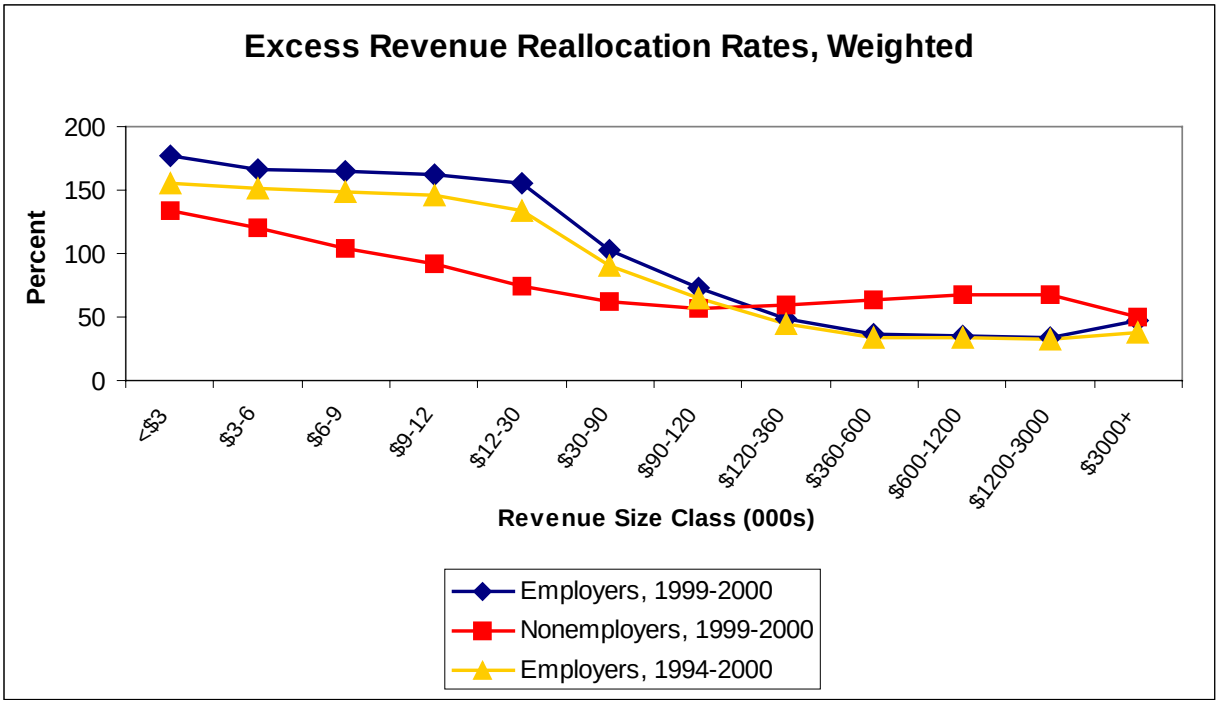


Figure 4.5 Mean Payroll Growth and Excess Payroll Reallocation by Employer Age